1. [25 points] Let $X$ be a random variable taking values in the interval $[0, 1]$, and let $Y$ be a random variable jointly distributed with $X$. We wish to estimate $X$ on the basis of $Y$.

(a) In preparation for part (b), prove the reverse Markov inequality: If $U$ is a random variable such that $P(U \leq c) = 1$ for some $c \in \mathbb{R}$, then for any $u < \mathbb{E}[U]$

$$P(U > u) \geq \frac{\mathbb{E}[U] - u}{c - u}$$

(Hint: $V = c - U$ is nonnegative with probability one.)

Solution: The event $\{U \leq u\}$ is equivalent to the event $\{V \geq c - u\}$. Since $V$ is nonnegative, Markov’s inequality gives

$$P(U \leq u) = P(V \geq c - u) \leq \frac{\mathbb{E}[V]}{c - u} = \frac{c - \mathbb{E}[U]}{c - u}$$

(note that both the numerator and the denominator are strictly positive because $u < \mathbb{E}[U] \leq c$). From this,

$$P(U > u) = 1 - P(U \leq u) \geq \frac{\mathbb{E}[U] - u}{c - u}.$$ 

(b) Let $e = X - \mathbb{E}[X|Y]$ be the error of the MMSE estimator of $X$ from $Y$. Suppose that the joint distribution of $X$ and $Y$ is such that $\mathbb{E}[e^2] = 1/2$. Use the reverse Markov inequality from part (a) to prove that

$$P\left(\mathbb{E}[e^2|X] > \frac{1}{4}\right) \geq \frac{1}{3}.$$

Solution: By the tower property of conditional expectations, $\mathbb{E}[\mathbb{E}[e^2|X]] = \mathbb{E}[e^2] = 1/2$. Since $X \in [0, 1]$, $\mathbb{E}[X|Y] \in [0, 1]$ as well, which means that $e^2 \in [0, 1]$. Applying the reverse Markov inequality with $U = \mathbb{E}[e^2|X]$, $c = 1$ and $u = 1/4$, we get

$$P\left(\mathbb{E}[e^2|X] > \frac{1}{4}\right) \geq \frac{\mathbb{E}[e^2] - 1/4}{1 - 1/4} = \frac{1/2 - 1/4}{1 - 1/4} = 1/3.$$ 

2. [25 points] A function $f : \mathbb{R} \to \mathbb{R}$ is $c$-Lipschitz for some $c > 0$ if $|f(x) - f(y)| \leq c|x - y|$ for all $x, y \in \mathbb{R}$.

(a) Let $\{X_n\}_{n=1}^{\infty}$ and $\{Z_n\}_{n=1}^{\infty}$ be two sequences of random variables, both converging in the mean square sense. Define a new sequence $\{Y_n\}_{n=1}^{\infty}$ of random variables by

$$Y_n = f(X_n) + Z_n, \quad n = 1, 2, \ldots$$

where $f$ is $c$-Lipschitz. Does the sequence $\{Y_n\}_{n=1}^{\infty}$ converge in the mean-square sense? Justify your answer. (You do not have to identify the limit.)
Solution: Yes, by the Cauchy criterion. For any \(m, n\) we have
\[
\mathbb{E}[(Y_m - Y_n)^2] = \mathbb{E}[(f(X_m) - f(X_n) + Z_m - Z_n)^2]
\]
\[
= \mathbb{E}[(f(X_m) - f(X_n))^2] + 2\mathbb{E}[(f(X_m) - f(X_n))(Z_m - Z_n)] + \mathbb{E}[(Z_m - Z_n)^2].
\]
Since \(f\) is \(c\)-Lipschitz, the first term on the right-hand side is bounded by
\[
c^2\mathbb{E}[(X_m - X_n)^2],
\]
which converges to zero as \(m, n \to \infty\) since \(X_n\) converges in mean square sense. The third term likewise converges to zero as \(m, n \to \infty\) since \(Z_n\) converges in mean square sense. For the middle term, use Cauchy–Schwarz and the Lipschitz property of \(f\):
\[
\mathbb{E}[(f(X_m) - f(X_n))(Z_m - Z_n)] \leq \sqrt{\mathbb{E}[(f(X_m) - f(X_n))^2] \mathbb{E}[(Z_m - Z_n)^2]}
\]
The same argument as before shows that both expectations under the square root converge to zero as \(m, n \to \infty\). Therefore, \(\mathbb{E}[(Y_m - Y_n)^2] \to 0\) as \(m, n \to \infty\), so the sequence \(\{Y_n\}\) converges in the mean square sense.

(b) Let \(\{X_n\}_{n=1}^\infty\) be a sequence of i.i.d. random variables taking values in the interval \([-1, 1]\), and for each \(n\) let \(Y_n = f(X_n)\), where \(f\) is a 1-Lipschitz function with \(\mathbb{E}[f(X_1)] = 0\). Give a direct proof (without using the Strong Law of Large Numbers) that the sequence
\[
\overline{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i, \quad n = 1, 2, \ldots
\]
converges almost surely, and identify the limit.

Solution: Each random variable \(Y_i = f(X_i)\) is bounded between \(\min_{x \in [-1,1]} f(x)\) and \(\max_{x \in [-1,1]} f(x)\). Since \(f\) is 1-Lipschitz,
\[
\max_{x \in [-1,1]} f(x) - \min_{x \in [-1,1]} f(x) \leq \max_{x, x' \in [-1,1]} |f(x) - f(x')| \leq 2.
\]
Moreover, each \(Y_i\) has zero mean. Using the Chernoff–Hoeffding bound, for any \(\epsilon > 0\) we get
\[
P\left(|\overline{Y}_n| \geq \epsilon\right) \leq P\left(\overline{Y}_n \geq \epsilon\right) + P\left(-\overline{Y}_n \geq \epsilon\right) \leq 2\exp\left(-\frac{n\epsilon^2}{2}\right).
\]
This implies that
\[
\sum_{n=1}^\infty P\left(|\overline{Y}_n| \geq \epsilon\right) \leq 2\sum_{n=1}^\infty \exp\left(-\frac{n\epsilon^2}{2}\right) < \infty, \quad \forall \epsilon \geq 0.
\]
Therefore, by the Borel–Cantelli lemma the sequence \(\{\overline{Y}_n\}\) converges to zero almost surely.

3. [25 points] Let \(X, Z_1\) and \(Z_2\) be three mutually independent \(N(0, 1)\) random variables. Let
\[
Y_1 = X + Z_1 + Z_2
Y_2 = X + 3Z_1 + Z_2
\]
(a) Find the simplest expression for $E[X|Y_1, Y_2]$

**Solution:** Because $X$, $Z_1$ and $Z_2$ are mutually independent Gaussian random variables, $X$, $Y_1$ and $Y_2$ are jointly Gaussian. Therefore, $E[X|Y_1, Y_2] = \hat{E}[X|Y_1, Y_2]$. Let $Y = (Y_1, Y_2)^T$. Then, because $X$ is zero mean,

$$\hat{E}[X|Y] = \text{Cov}(X, Y)\text{Cov}(Y)^{-1}Y,$$

assuming that $\text{Cov}(Y)$ is nonsingular. Using the fact that $X, Y_1, Y_2$ are independent $N(0, 1)$ random variables, we have

$$\text{Cov}(X, Y) = (E[XY_1] E[XY_2]) = (E[X^2] E[X^2]) = (1 1)$$

$$\text{Cov}(Y) = \begin{pmatrix} E[Y_1^2] & E[Y_1 Y_2] \\ E[Y_2 Y_1] & E[Y_2^2] \end{pmatrix}$$

$$= \begin{pmatrix} E[(X + Z_1 + Z_2)^2] & E[(X + Z_1 + Z_2)(X + 3Z_1 + Z_2)] \\ E[(X + 3Z_1 + Z_2)(X + Z_1 + Z_2)] & E[(X + 3Z_1 + Z_2)^2] \end{pmatrix}$$

$$= \begin{pmatrix} E[X^2 + Z_1^2 + Z_2^2] & E[X^2 + 3Z_1^2 + Z_2^2] \\ E[X^2 + 3Z_1^2 + Z_2^2] & E[X^2 + 9Z_1^2 + Z_2^2] \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 5 \\ 5 & 11 \end{pmatrix}.$$

Indeed, $\det \text{Cov}(Y) = 8$. Therefore,

$$\text{Cov}(Y)^{-1} = \frac{1}{\det \text{Cov}(Y)} \begin{pmatrix} 11 & -5 \\ -5 & 3 \end{pmatrix} = \frac{1}{8} \begin{pmatrix} 11 & -5 \\ -5 & 3 \end{pmatrix}$$

and

$$E[X|Y_1, Y_2] = \hat{E}[X|Y_1, Y_2] = \text{Cov}(X, Y)\text{Cov}(Y)^{-1}Y$$

$$= \frac{1}{8} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 11 & -5 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

$$= \frac{3Y_1}{4} - \frac{Y_2}{4}.$$

(b) Find the simplest expression for the corresponding MSE for estimation of $X$ by $E[X|Y_1, Y_2]$.

**Solution:** From the solution to part (a),

$$X - E[X|Y_1, Y_2] = X - \frac{3Y_1}{4} + \frac{Y_2}{4}$$

$$= X - \frac{3}{4}(X + Z_1 + Z_2) + \frac{1}{4}(X + 3Z_1 + Z_2)$$

$$= X - \frac{Z_2}{2}$$

Therefore, $\text{MSE} = \frac{1}{4}E[(X - Z_2)^2] = \frac{1}{2}$.
4. [25 points] Let $X$ and $Y$ be jointly Gaussian random variables with zero mean, such that the vector $\begin{pmatrix} X \\ Y \end{pmatrix}$ has covariance matrix $\begin{pmatrix} 4 & 4 \\ 4 & 8 \end{pmatrix}$.

(a) Find the conditional expectation $\mathbb{E}[e^tX | Y]$. (Your answer should be a function of $Y$, and it should depend on $t$ as well.)

**Solution:** For any $y$, the conditional distribution of $X$ given $Y = y$ is $N(\mathbb{E}[X|Y = y], \text{Var}(e))$, where $e = X - \mathbb{E}[X|Y] = X - \bar{E}[X|Y]$. We have

$$\bar{E}[X|Y] = \frac{\text{Cov}(X,Y)Y}{\text{Var}(Y)} = \frac{4Y}{8} = \frac{Y}{2}$$

$$\text{Var}(e) = \text{Var} \left( X - \frac{Y}{2} \right)$$

$$= \text{Var}(X) - \frac{1}{2} \text{Cov}(X,Y) - \frac{1}{2} \text{Cov}(Y,X) + \frac{1}{4} \text{Var}(Y)$$

$$= \text{Var}(X) - \text{Cov}(X,Y) + \frac{1}{4} \text{Var}(Y) = 2.$$

Therefore, $\mathbb{E}[e^tX | Y = y]$ is equal to the moment generating function of a $N(y/2, 2)$ random variable, which is $e^{ty/2 + t^2}$. The answer is $\mathbb{E}[e^tX | Y] = e^{ty/2 + t^2}$ (note that this is a random variable).

(b) Express the conditional probability $P(|X| \geq 4 | Y)$ in terms of the standard Gaussian cdf $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-x^2/2} dx$. (Your answer should be a function of $Y$.)

**Solution:** From part (a), the conditional distribution of $X$ given $Y = y$ is $N(y/2, 2)$. If $Z$ has a standard Gaussian distribution $N(0, 1)$, then

$$P(|X| \geq 4 | Y = y) = P(\sqrt{2}Z + y/2 \geq 4)$$

$$= P(\sqrt{2}Z \geq 4 - y/2) + P(\sqrt{2}Z \leq -4 - y/2)$$

$$= P \left( Z \geq \frac{4 - y/2}{\sqrt{2}} \right) + P \left( Z \leq \frac{-4 - y/2}{\sqrt{2}} \right)$$

$$= 1 - \Phi \left( \frac{4 - y/2}{\sqrt{2}} \right) + \Phi \left( \frac{-4 - y/2}{\sqrt{2}} \right)$$

Consequently,

$$P(|X| \geq 4 | Y) = 1 - \Phi \left( \frac{4 - Y/2}{\sqrt{2}} \right) + \Phi \left( \frac{-4 - Y/2}{\sqrt{2}} \right)$$

(note that this is a random variable).